



# Arquitetura de Plataforma de Dados e Modern Data Stack

## Componentes-Chave da Arquitetura de Dados



## Arquitetura de Plataforma de Dados e Modern Data Stack

---

A arquitetura de dados é uma estrutura projetada para garantir que os dados sejam coletados, armazenados, gerenciados, integrados e entregues aos usuários de maneira otimizada. Os componentes-chave da arquitetura de dados incluem:

**Modelo de Dados:** Fornece uma representação visual dos dados e como eles estão relacionados. Os modelos podem ser lógicos, mostrando como os dados são percebidos pelo usuário final, ou físicos, mostrando como os dados são organizados internamente.

**Bancos de Dados:** Neste componente temos diferentes categorias. A escolha depende do tipo de dado que a empresa armazena e mais uma escolha é normalmente o mais comum.

- Relacionais: Como MySQL, PostgreSQL, SQL Server, Oracle.
- NoSQL: Como MongoDB, Cassandra, CouchDB.
- Armazéns de Dados (Data Warehouses): Como Redshift, Teradata, Snowflake.
- Bases de Dados Temporais: Focadas em dados de séries temporais, como InfluxDB.

**Integração de Dados:** Aqui também teremos diferentes categorias dependendo dos dados usados pela empresa.

- ETL (Extração, Transformação e Carga): Ferramentas como Apache NiFi, Talend, Informatica.
- Streaming de Dados: Kafka, Apache Pulsar, RabbitMQ.
- Lago de Dados (Data Lake): Repositório centralizado que permite armazenar todos os dados estruturados e não estruturados em sua forma original. Exemplos incluem Amazon S3, Hadoop HDFS, Azure Data Lake.

**Qualidade dos Dados:** Ferramentas e processos que ajudam a garantir a precisão, completude, atualidade e consistência dos dados.

**Metadados:** Informações sobre os dados, incluindo descrições, origens, relacionamentos, padrões e outras características.

**Governança de Dados:** Conjunto de políticas, procedimentos, estrutura organizacional e padrões para gerenciar a integridade, qualidade, segurança e uso dos dados.

**Segurança de Dados:** Proteção dos dados contra acessos não autorizados, incluindo criptografia, controle de acesso e auditorias.

**Soluções de Big Data:** Plataformas como Hadoop e Spark para processar e analisar grandes volumes de dados.

## ***Arquitetura de Plataforma de Dados e Modern Data Stack***

---

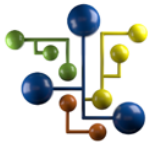
**Análise e BI (Business Intelligence):** Ferramentas que permitem a análise de dados para tomada de decisão, como Tableau, QlikView, Power BI.

**APIs e Serviços de Dados:** Interface para permitir que sistemas e aplicações externas acessem os dados de maneira controlada e segura.

**Arquitetura em Camadas:** Divisão da arquitetura em zonas ou camadas, como zona de landing, zona de processamento e zona de consumo.

**Gerenciamento do Ciclo de Vida dos Dados (DLM):** Processos e ferramentas para gerenciar os dados durante todo o seu ciclo de vida, desde a criação até a eliminação.

Estes são apenas os principais componentes e, dependendo da necessidade da organização, da complexidade e do volume dos dados, outros componentes específicos podem ser adicionados à arquitetura.



**Equipe DSA**

Muito Obrigado!  
Continue Trilhando Uma Excelente Jornada de Aprendizagem.